DOI:
10.1109/TBME.2018.2865669

Document Version
Peer reviewed version

Regional Multi-view Learning for Cardiac Motion Analysis: Application to Identification of Dilated Cardiomyopathy Patients

Esther Puyol-Antón*, Bram Ruijsink*†, Bernhard Gerber‡, Mihaela Silvia Amzulescu‡,
Hélène Langet§, Mathieu De Craene§, Julia A. Schnabel§, Paolo Piro§, and Andrew P. King§
*School of Biomedical Engineering & Imaging Sciences, King's College London, U.K
†Guy's and St Thomas' Hospital NHS Foundation Trust, London, UK.
‡Philips Research, Medisys, Paris, France.
§Division of Cardiology, Cliniques Universitaires St-Luc, Avenue Hippocrate 10-2881, B-1200 Brussels, Belgium.

Abstract*—*Objective: **The aim of this paper is to describe an automated diagnostic pipeline which uses as input only ultrasound (US) data, but is at the same time informed by a training database of multimodal magnetic resonance (MR) and US image data.** ***Methods:*** **We create a multimodal cardiac motion atlas from 3D MR and 3D US data followed by multi-view machine learning algorithms to combine and extract the most meaningful cardiac descriptors for classification of dilated cardiomyopathy (DCM) patients using US data only. More specifically, we propose two algorithms based on: multi-view linear discriminant analysis (MLDA) and multi-view Laplacian support vector machines (MvLapSVM). Furthermore, a novel regional multi-view approach is proposed to exploit the regional relationships between the two modalities.** ***Results:*** **We evaluate our pipeline on the classification task of discriminating between normals and DCM patients. Results show that the use of multi-view classifiers together with a cardiac motion atlas results in a statistically significant improvement in accuracy compared to classification without the multimodal atlas. MvLapSVM was able to achieve the highest accuracy for both the global approach (92.71%) and the regional approach (94.32%).** ***Conclusion:*** **Our work represents an important contribution to the understanding of cardiac motion, which is an important aid in the quantification of the contractility and function of the left ventricular myocardium.** ***Significance:*** **The intended workflow of the developed pipeline is to make use of the prior knowledge from the multimodal atlas to enable robust extraction of indicators from 3D US images for detecting DCM patients.**

***Index Terms*—Cardiac motion atlas, multi-modality, multi-view classification**

I. Introduction

Assessment of left ventricular (LV) function is important in the diagnosis of cardiovascular disease. LV function has been traditionally assessed using global indicators such as ejection fraction, stroke volume or ventricular mass. The main drawback of such indicators is that they do not provide localised information about ventricular function, which limits their diagnostic power. More localised indicators are sometimes based on the 17-segment model proposed by the American Heart Association (AHA) [1]. This divides the image data into regional segments using criteria based on anatomical landmarks. More specifically, the recommendation given by the AHA is to divide the left ventricle of the heart into 17 regions with six regions for the basal area (1-6), six regions for the mid area (7-12) and five for the apical areas (13-17). However, the AHA regions have coarse correspondence with underlying coronary perfusion territories. Myocardial wall motion analysis (MWMA) offers a much more localised and flexible description of LV function. Using MWMA, myocardial wall motion can be characterised from imaging data using displacement and velocity, and myocardial deformation can be expressed as strain or strain ratio. This allows for a more spatially and temporally localised assessment of LV function.

3D tagged magnetic resonance (MR) imaging is increasingly accepted as the gold standard for MWMA due to its excellent image contrast and coverage [2]. However, its clinical use is limited by lack of access to scanners, high acquisition cost and a lack of expertise in acquiring the data. Instead, ultrasound (US) is commonly used in the clinic due to its low acquisition cost and portability. However, coverage of the heart using US is restricted due to limited acoustic windows.

In Puyol-Antón *et al.* [3], we proposed a multimodal cardiac motion atlas that was able to relate the MWMA parameters derived from MR and US data. This opened up the possibility of exploiting this relationship to enable a diagnostic pipeline based on US data that at the same time is informed by a training database of multimodal (MR and US) data. This is the focus of this paper, and we achieve this aim by proposing a number of novel extensions to the multimodal atlas. We demonstrate the use of our new approach on the identification of dilated cardiomyopathy (DCM) patients using only US data. In the following sections, we first present an overview of motion atlases in the literature, followed by relevant work from the field of multi-view machine learning, which our novel extensions are based on. Finally we review the literature on DCM patients, and summarise our novel contributions in this context.

A. Cardiac motion atlases

Cardiac motion atlases provide a space of reference in which population comparisons of motion (i.e. displacements and velocities) and deformation (i.e. strains and strain rates) can be carried out. The use of such atlases for the statistical analysis of normal and pathological LV function has gained increasing interest over the past decade. The main steps involved in the formation of a cardiac motion atlas are: definition of the geometry, spatial and temporal normalisation, and reorientation of subjects' cardiac geometry and motion, both spatially and over time. These steps transform the subject-specific motion/deformation data so that they can be directly compared, removing biases due to heart orientation, size, shape and cardiac phase.

This work is funded by the King's College London & Imperial College London EPSRC Centre for Doctoral Training in Medical Imaging (EP/L015226/1) and EPSRC grant EP/R005516/1. This work was supported by an EPSRC programme Grant (EP/P001009/1) and the Wellcome EPSRC Centre for Medical Engineering at Kings College London (WT 203148/Z/16/Z) and by the National Institute for Health Research (NIHR) Biomedical Research Centre award to Guy and St Thomas NHS Foundation Trust in partnership with King's College London, and by the NIHR Healthcare Technology Cooperative for Cardiovascular Disease at Guys and St Thomas NHS Foundation Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Typically, cardiac motion atlases have focused on the LV, as it is the chamber primarily investigated for diagnosing cardiac diseases. The motion/deformation estimates have been made from imaging data acquired using MR [4], [5], [6], [7], US [8], [9], both MR and US [3] or Computed Tomography (CT) [10]. Applications of cardiac motion atlases have included the statistical analysis of normal and pathological LV motion [8], [9], [7], the identification of disease [4], [11] and prospectively predicting treatment outcome [8], [6], [5].

B. Multi-view machine learning

The incorporation of multimodal (MR and US) data into a cardiac motion atlas as described in Puyol-Antón *et al.* [3] raises the question of how best to exploit these data for diagnosis. In machine learning, the field of multi-view learning has developed a range of techniques for dealing with situations in which there are multiple 'views' of the same underlying phenomenon. We propose to consider MR and US as two ways of viewing the mechanics of the heart, and this perspective enables us to exploit the rich literature from the field of multi-view machine learning.

A naive solution for multi-view learning is to concatenate the data of the multiple views and apply single-view learning algorithms directly. However, one problem with this approach is over-fitting, as the number of features would increase dramatically. Another problem is that it would not take advantage of specific statistical properties of each view; different views often contain complementary information, and multi-view learning can exploit this information to learn a joint representation that is more expressive than that of single-view learning methods. To overcome these problems, several more sophisticated multi-view machine learning algorithms have been proposed to improve the generalisation performance. We focus in particular on multi-view dimensionality reduction and supervised learning algorithms [12].

Multi-view dimensionality reduction algorithms aim to construct linear or non-linear transformations from the original high-dimensional spaces of the different views to a new space, under the constraint that the multiple transformed feature sets should be aligned in space. When correspondences between the data from the two views are known, most of these algorithms are based on canonical correlation analysis (CCA) [13], and its non-linear extension, Kernel CCA [14]. Multi-view dimensionality reduction algorithms are also often referred as manifold alignment algorithms, although in this case at least some of the inter-modality correspondences are typically unknown.

In multi-view supervised learning, a typical approach is to add a regularisation term to the objective function of the dimensionality reduction to constrain intra-class and inter-class characteristics. These approaches can be based on linear discriminant analysis (LDA). Some examples are multi-view LDA [15] and multi-view discriminant analysis [16].

Multi-view machine learning algorithms are trained using both views, but in some cases, they can be applied using only the information from one view, taking advantage of the learnt information for the required task. In this work, we take advantage of this characteristic to perform classification using motion data derived from only one view (i.e. US data).

Multiple kernel learning algorithms (MKL) have also been proposed to combine data from multiple views, or different feature mappings of the same view. MKL finds the optimal linear or non-linear combination of multiple kernels [17], addressing issues related to differences in representation, variability and dimensionality. However, MKL is not specifically intended to combine information from multiple views. Rather, it works by simply assigning weights to the kernels [17]. MKL has been applied in medical imaging. For example, in the context of cardiac imaging, Peressutti *et al.* [6] used MKL to combine motion data extracted using a cardiac motion atlas and non-imaging data. Also, Sanchez-Martinez *et al.* [18] proposed to use MKL to jointly analyse the variability of multiple velocity patterns from a stress protocol to improve the characterisation of heart failure with preserved ejection fraction patients. Finally, Marciniak *et al.* [19] proposed a MKL-based framework to classify between patients with ventricular fibrillation and patients presenting with other non-arrhythmic symptoms when they are hospitalised.

Of the algorithms discussed above, only MKL algorithms have been used with cardiac motion atlases, and only for combining different types of data, not for combining data from different imaging modalities. In this paper we aim to integrate and find relationships between data from two modalities (MR and US). In Sections III-B and III-C we describe how we exploit and extend multi-view learning methods for use with our multimodal motion atlas, with the aim of relating the MWMA parameters derived from MR and US data.

C. Dilated cardiomyopathy

In this work we develop and evaluate our novel method on the problem of identifying DCM patients. Therefore, we include here a brief review of DCM.

Cardiomyopathies are a group of myocardial disorders in which structural and functional abnormalities of the heart muscle develop in the absence of clear cardiovascular causes, such as hypertension, valvular disease, coronary artery disease or congenital abnormalities. DCM is the most common cardiomyopathy. The global myocardial dysfunction is progressive and often irreversible, resulting in heart failure (HF) in the majority of patients [20]. The incidence of DCM is estimated to be between 5 and 8 per 100,000 people and is the third most common cause of HF [21]. Furthermore, due to the severity and irreversibility of the disease, DCM is one of the most common indicators for heart transplantation.

Identification of DCM relies primarily upon imaging data such as US or MR, as they allow accurate quantification of LV function, and identification and localisation of scar. The diagnostic criteria for DCM are [20]: reduced global ejection fraction, signs of high LV filling pressure and LV or biventricular enlargement, and depressed myocardial performance and global contractile dysfunction of one or both ventricles. Features of regional wall motion abnormalities, such as mechanical dyssynchrony and reduced regional myocardial strain, are known to have additional beneficial value for improved disease classification and treatment stratification. However, such tools are still underused due to the labour intensity of currently available methods [22].

The main limitation of the current apporaches for DCM diagnosis is the variation of etiologies of the disease. Actually, more than 75 known conditions can present the DCM phenotype [23]. Due to the multi-factorial development of the disease, determination of its origin and thus appropriate treatment for each patient is a challenging task. Furthermore, DCM patients have reported heterogeneous regional values for myocardial perfusion, systolic function, asynchrony and myocardial work. A longer term aim of this work is to find new local descriptors that can help to further stratify DCM patients.

D. Our contributions

In this paper we build on our previous work [3], in which we developed a multimodal spatiotemporal cardiac motion atlas from MR and US data. The novelties we present in this work are twofold: (1) We perform multi-view dimensionality reduction and classification in a single step, enabling the embedding to be optimised for the

classification task; (2) we introduce a novel regional approach for multi-view classification, which enables the regional dependence of the multi-view relationship to be exploited in the classification task. Preliminary work (using only a linear global multi-view classifier) was presented in Puyol-Antón *et al.* [24]. To the authors' knowledge, this is the first time that multi-view machine learning has been used for disease diagnosis using multimodal data.

The remainder of this paper is organised as follows. In Section II, we describe details of the clinical data sets used for evaluation. In Section III we describe the framework that we have developed to build and apply the multimodal spatiotemporal cardiac motion atlas, including a description of the different multi-view learning algorithms employed. Results are presented in Section IV, while Section V discusses the findings of this paper in the context of the literature and proposes potential improvements for future work.

II. Materials

A. Multimodal data set

Three clinical MR and 3D US data sets were used for evaluation in this paper. The first is the database used for the cardiac motion analysis challenge that was held at the 2011 MICCAI workshop "Statistical Atlases and Computational Models of the Heart: Imaging and Modelling Challenges" (STACOM'11) [25]. The STACOM'11 database includes MR and US data from 15 healthy volunteers acquired at the Biomedical Engineering & Imaging Sciences, King's College London, United Kingdom, and the Department of Internal Medicine II - Cardiology, University of Ulm, Germany.

The second data set was acquired at the Division of Cardiology, Cliniques Universitaires St-Luc, Avenue Hippocrate 10-2881, B-1200 Brussels, Belgium. This contains MR and US data acquired from 26 healthy volunteers and 19 patients with DCM.

The third data set was more recently acquired at the School of Biomedical Engineering & Imaging Sciences, King's College London, United Kingdom, and contains MR and US data acquired from a further 9 healthy volunteers.

The first and second MR data sets were acquired using a 3T Philips Achieva System (Philips Healthcare, Best, The Netherlands), while the third data set was acquired using a 1.5T Philips Ingenia System (Philips Healthcare, Best, The Netherlands). The US data sets were all acquired using an iE33 3D echocardiography system (Philips Medical Systems, Bothell, WA, United States) with a 1-5 MHz transthoracic matrix array transducer (xMATRIX X5.1). Full-volume acquisition mode was used in which several smaller imaging sectors acquired over multiple cardiac cycles are combined to form a large composite volume.

In particular, the three data sets contain for each subject:

- **cine SA and LA**: a multi-slice short-axis (SA) cine-MR sequence covering the full heart, and two orthogonal long-axis (LA) planes (2-chamber (2Ch) and 4-chamber (4Ch) views) (TR/ TE = $3.0/1.5$ ms, flip angle = $60°$). Typical slice thickness was between 8.0 - 10.0 mm for SA and LA sequences respectively, with an in-plane resolution between 1.0 - 1.4 mm $\times$ 1.0 - 1.4 mm. Typical temporal resolution was between 25 and 30 frames per cycle. In our framework, only the ED frames of the cine SA and LA data were used for geometry estimation (see Section III-A). The cine SA and LA sequences were not used for motion estimation.
- **TAG**: a 3D tagged MR sequence in three orthogonal directions (TR/ TE = $7.5/3.2$ ms, flip angle = $19°$, tag distance between 7.7 -8.8 mm). The images have reduced field-of-view enclosing the LV, with typical isotropic 3D spatial resolution between 2.5 mm and 1.1 mm, and typical temporal resolution between 22 and 30 frames per cycle. The TAG data were used for motion estimation (see Section III-A).
- **US**: an apical 3D LV full-volume ultrasound sequence. Typical image resolution and size varied from 0.7 mm - 1.0 mm and $208 \times 240 \times 176$ to $272 \times 256 \times 224$ voxels, respectively. Typical temporal resolution was between 15 and 23 frames per cycle. The US data were used for both geometry and motion estimation (see Section III-A).

The three data sets described above were combined to form a single data set for the experiments described in this paper. The total amount of included cases was: 19 patients with DCM and 50 healthy volunteers.

Demographics of the entire cohort are shown in Table I. LV volume was computed from manual endocardial segmentations using Simpson's rule. A Student's t-test (99% confidence with unequal variances) was used to compare the volume values between groups and modalities.

		Healthy volunteers	DCM patients
Study population, n		50	19
Age (years)		37 (20-77)	54 (30-79)
MR	EDV (mL)	169.2 (24.8)*†	262.9 (79.3)
	ESV (mL)	72.54 (14.1)*†	193.26 (74.7)
	EF (%)	56.9 (7.2)*†	27.9 (12.3)
	LV mass (g)	101.7 (14.7)*†	73.2 (30.5)
US	EDV (mL)	128.5 (14.7)*	237.7 (74.8)
	ESV (mL)	61.1 (16.8)*	174.2(72.6)
	EF (%)	52.07 (8.9)*	24.1 (11.8)
	LV mass (g)	92.9 (19.2)*	66.8 (26.6)

Table I: Study demographics: end-diastolic volume (EDV), end-systolic volume (ESV), ejection fraction (EF) and LV mass. All expressed as mean (standard deviation) and age expressed as mean (min-max). An asterisk indicates a statistically significant difference between DCM and healthy volunteers, and a dagger indicated a statistically significant difference between modalities.

Fig. 1 shows an example of LV MR and US acquisitions for a healthy volunteer, and a patient suffering from DCM.

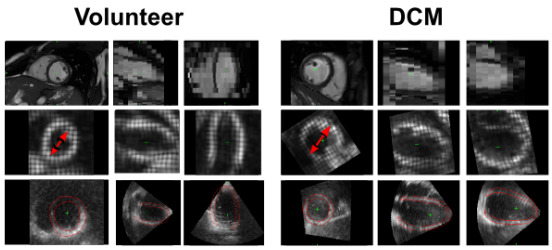

Figure 1: Left: An example of MR and 3D US acquisitions for a healthy volunteer (top: cine MR, middle: tagged MR, bottom: US). Right: for a patient suffering from DCM. Red arrows indicates the dilation of the LV in patients suffering from DCM compared to the healthy volunteer.

III. Methods

Previously, in Puyol-Antón *et al.* [3] we proposed a pipeline to form a multimodal spatiotemporal 3D cardiac motion atlas of the LV from MR and US data. This atlas allows direct comparison of motion parameters estimated from different subjects by removing differences in shape and cardiac cycle duration from the comparison. The main novelty of the work we present here lies in developing a diagnostic pipeline using only 3D US data, but at the same time taking advantage of the implicit relationship between MR and US learned

from the atlas. This framework may provide a better understanding of the mechanisms responsible for the development and progression of cardiac disease which could assist therapeutic planning and improve treatments. As an exemplar application this paper illustrates the use of the framework for classifying subjects as DCM/normal.

In the following, Section III-A briefly reviews the main steps involved in forming the multimodal motion atlas as outlined in Puyol-Antón *et al.* [3]; Section III-B introduces our novel global approach to classify patients with DCM using multi-view machine learning; finally Section III-C presents our novel extension of this multi-view learning approach to better capture regional variations in the motion parameters. The proposed framework is illustrated in Fig. 2.

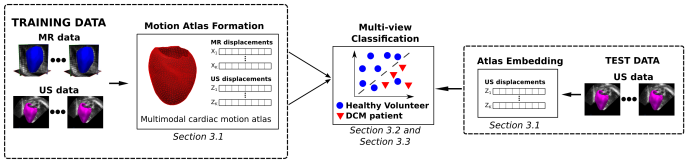

Figure 2: Overview of the proposed framework for forming and applying a multimodal motion atlas.

A. Motion atlas formation

The following sections are a brief summary of the atlas formation process, included here for completeness. Figure 3 shows the main steps involved in the atlas formation. For further details, please refer to Puyol-Antón *et al.* [3].

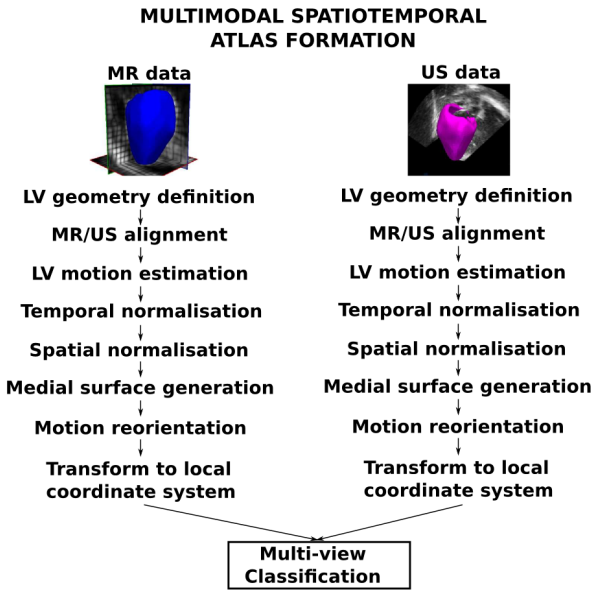

Figure 3: Overview of the framework for spatiotemporal cardiac motion atlas formation (figure adapted from Puyol-Antón *et al.* [3]).

LV geometry definition: For each subject, cine SA/LA and US images were manually segmented in the end-diastolic (ED) phase. SA and LA segmentations were used to correct for breath-hold induced motion artefacts using the iterative registration algorithm proposed in Sinclair *et al.* [26]. The motion-corrected LA/SA segmentations were fused to form a smooth myocardial mask, and then a statistical shape model (SSM) was fitted to the mask to generate a surface mesh with point correspondence for each of the subjects and 17 AHA regions.

Medial surface generation: A medial surface mesh with regularly sampled vertices ($\approx$ 1000) was generated from the personalised SSM epicardial and endocardial mesh. The use of a medial surface enables a more robust motion estimation compared to the endo- and epicardial surfaces of the SSM as it is likely to be less affected by motion tracking errors which can be caused by inaccuracies in the ED segmentation.

MR/US alignment: MR and US images were aligned by registering their respective LV meshes using a Generalised Procrustes analysis [27]. AHA delineations in both meshes were used to ensure that the midseptum was in the same location in both modalities.

LV motion estimation: A B-spline free-form deformation (FFD) registration was used [28] to estimate LV motion between consecutive frames of the MR and US sequences. A 3D+*t* cyclic B-spline was fitted to the composed 3D transformations in order to estimate a full cycle 3D+*t* transformation [29].

Temporal normalisation: This stage establishes temporal correspondence between all subjects based on specific cardiac events and transforms them to a normalised timescale. Three cardiac events were automatically identified from the volume curves and used in a piecewise linear temporal transformation. Note that this normalisation is applied separately to the MR and US data, but both are transformed to the same normalised timescale. The piecewise linear temporal transformation, together with the 3D+*t* cyclic transformation, was used to propagate the medial surface mesh to a common set of time points for both modalities and all subjects. Note that all LV displacements were represented as differences between mesh locations in the current frame and the ED frame.

Spatial normalisation: This step aims to remove bias towards differences in subject-specific LV geometries from the motion analysis. Similar to Peressutti *et al.* [6] and Puyol-Antón *et al.* [3], each mesh is transformed to an unbiased atlas coordinate system using a combination of Procrustes alignment and Thin Plate Spline transformation.

Motion reorientation: For each subject, LV displacements derived from MR and US for each vertex and each cardiac phase were reoriented into the atlas coordinate system under a small deformation assumption using a push-forward action [8], [30].

Transform to local coordinate system: For a more intuitive understanding of the LV motion, MR displacements in the atlas coordinate system were projected onto a local cylindrical coordinate system, providing radial, longitudinal and circumferential information. The longitudinal direction is defined uniformly by drawing a line from the apex to the mitral valve. The same projection was applied to the US-derived displacements.

Remove vertices outside the FoV: For each subject, the vertices of the reference ED medial mesh that fell outside of the field-of-view (FoV) of the MR or US images were removed from the analysis. The intersection of all subjects' FoVs was computed to only retain the points that fell inside the FoV for all subjects. Most of the vertices removed in this way belonged to the LV apex.

B. Multi-view classification: A global approach

We now describe our novel methodological contributions. In this section we present our application of multi-view machine learning techniques to exploit the information in the multimodal atlas for a supervised classification task.

The local MR displacements in atlas space were concatenated into a row vector such that for subject k, $\hat{\boldsymbol{x}}_k \in \mathbb{R}^{1\times D}$, where D = (3 $\times$ N $\times$ M) with N the number of cardiac phases and M the number of points in the atlas medial surface mesh. The row vectors $\hat{\boldsymbol{x}}_k$ for each subject were stacked to produce a matrix $X = [\hat{\boldsymbol{x}}_1^{\mathrm{T}}, ..., \hat{\boldsymbol{x}}_K^{\mathrm{T}}] \in \mathbb{R}^{K\times D}$, where K is the number of subjects. Likewise, the US displacements in atlas space were concatenated to form the matrix $Z = [\hat{\boldsymbol{z}}_1^{\mathrm{T}}, ..., \hat{\boldsymbol{z}}_K^{\mathrm{T}}] \in \mathbb{R}^{K\times D}$.

We aim to classify subjects as DCM/normal using only US data, but taking advantage of the multimodal cardiac motion atlas. To this end, we propose to apply a multi-view classifier, which reduces the dimensionality and performs the classification simultaneously. We evaluate the use of linear and non-linear multi-view classifiers, the details of which are discussed below.

Linear method: For the linear case, we used the multi-view linear discriminant analysis (MLDA) [15] algorithm, which seeks to find a common space while simultaneously preserving the correlation between modalities and the discriminating information in each modality. The optimisation problem of MLDA is given by:

$$\begin{aligned} \max_{\boldsymbol{w_x}, \boldsymbol{w_z}} \quad & \boldsymbol{w_x}^{\mathrm{T}} S_{bx} \boldsymbol{w_x} + \boldsymbol{w_z}^{\mathrm{T}} S_{bz} \boldsymbol{w_z} + 2\gamma \boldsymbol{w_x}^{\mathrm{T}} C_{xz} \boldsymbol{w_z} \\ \text{subject to} \quad & \boldsymbol{w_x}^{\mathrm{T}} C_{xx} \boldsymbol{w_x} + \sigma \boldsymbol{w_z}^{\mathrm{T}} C_{zz} \boldsymbol{w_z} = 1 \end{aligned} \tag{1}$$

where $\boldsymbol{w_x}$ is the projection matrix for the first view (i.e. MR) and $\boldsymbol{w_z}$ is the projection matrix for the second view (i.e. US); S_{bx} and S_{bz} denote the between-class matrices for each view [15]; C_{xx} and C_{zz} are respectively the covariance matrices of the first view X (i.e. MR) and the second view Z (i.e. US); C_{xz} is the cross-covariance matrix of the two views (i.e. MR/US); and $\sigma = tr(C_{xx})/tr(C_{zz})$. The leftmost two terms of Eq. 1 attempt to minimise the within-class distance, while the rightmost term attempts to find a common space between the two views. γ is a regularisation parameter that balances the relative significance between these two objectives [15]. Note that Eq. 1 has a closed form solution.

In Eq. 1, the correlation between the different views and the discrimination of each view can be maximised simultaneously. Therefore, using the Lagrangian multiplier technique, the equation can be solved by a generalised multivariate eigenvalue problem as follows [15]:

$$\begin{bmatrix} S_{bx} & \gamma C_{xz} \\ \gamma C_{xz}^T & S_{bz} \end{bmatrix} \begin{bmatrix} \boldsymbol{w_x} \\ \boldsymbol{w_x} \end{bmatrix} = \lambda \begin{bmatrix} C_{xx} & 0 \\ 0 & \sigma C_{zz} \end{bmatrix} \begin{bmatrix} \lambda_x \boldsymbol{w_x} \\ \lambda_z \boldsymbol{w_z} \end{bmatrix} \tag{2}$$

After solving the generalised eigenvalue problem of Eq. 2 and retaining only the first d components, a subject k can be embedded into the low dimensional space according to the following strategies [15]:

- Using only MR data: $\boldsymbol{u}_k = \boldsymbol{w_x}^T \hat{\boldsymbol{x}}_k$
- Using only US data: $\boldsymbol{u}_k = \boldsymbol{w_z}^T \hat{\boldsymbol{z}}_k$
- Using both MR and US data: $\boldsymbol{u}_k = \boldsymbol{w_x}^T \hat{\boldsymbol{x}}_k + \boldsymbol{w_z}^T \hat{\boldsymbol{z}}_k$

In the low dimensional space, a n-nearest neighbours algorithm is used to classify between healthy and DCM patients, and produce the predicted label, $\hat{y_k}$.

Non-linear method: For the non-linear case, we modified the semi-supervised multi-view Laplacian support vector machines (MvLapSVM) algorithm introduced in Sun *et al.* [31] to only use labelled data. MvLapSVM integrates manifold regularisation, which imposes local constraints on samples, and multi-view learning. MvLapSVM is an extension of the Laplacian SVM algorithm for multi-view data. The MvLapSVM primal problem can be formulated as:

$$\begin{aligned} \min_{\boldsymbol{\alpha_x}, \boldsymbol{\alpha_z}, \boldsymbol{\xi_x}, \boldsymbol{\xi_z}} \quad & \frac{1}{2K} \sum_{i=1}^{K} (\xi_x^i + \xi_z^i) + \gamma_1 (\boldsymbol{\alpha_x}^{\mathrm{T}} G_x \boldsymbol{\alpha_x} + \boldsymbol{\alpha_z}^{\mathrm{T}} G_z \boldsymbol{\alpha_z}) + \\ & \gamma_2 (\boldsymbol{\alpha_x}^{\mathrm{T}} G_x L_x G_x \boldsymbol{\alpha_x} + \boldsymbol{\alpha_z}^{\mathrm{T}} G_z L_z G_z \boldsymbol{\alpha_z}) + \\ & \gamma_3 (G_x \boldsymbol{\alpha_x} - G_z \boldsymbol{\alpha_z})^{\mathrm{T}} (G_x \boldsymbol{\alpha_x} - G_z \boldsymbol{\alpha_z}) \\ \text{subject to} \quad & y_i \left(\sum_{j=1}^{K} \alpha_x^j g_x(x_j; x_i) \right) \geq 1 - \xi_x \\ & y_i \left(\sum_{j=1}^{K} \alpha_z^j g_z(z_j; z_i) \right) \geq 1 - \xi_z \\ & \xi_x^i,\, \xi_z^i \geq 0, \qquad i = 1, ..., K \end{aligned} \tag{3}$$

where $\boldsymbol{\alpha_x}, \boldsymbol{\alpha_z}$ are respectively the weight vectors for the first view X (i.e. MR) and the second view Z (i.e. US); $\boldsymbol{\xi_x}$ and $\boldsymbol{\xi_z}$ are the slack variables for the first view and the second view, used to make the objective function differentiable; G_x and G_z are the Gram matrices for the first and second views; L_x and L_z are the graph Laplacian matrices for the first and the second views; $\gamma_1, \gamma_2, \gamma_3 > 0$ are the norm regularisation terms; $y_i \in \{-1, 1\}^K$ are the classification labels and $k(\cdot, \cdot)$ is the kernel function.

The leftmost three terms of Eq. 3 can be rewritten as the sum of two Laplacian SVM optimisers [32], which assumes that the marginal probability distribution underlying the data is supported on a low dimensional manifold, and the Laplacian matrix learning exploits the connections among features, which is able to generalise over intra-class variations and provide inter-class discrimination. The rightmost term corresponding to γ_3 models the multi-view learning, using the assumption that a good learner can learn from each view, and consequently, these good learners in different views should be consistent to a large extent with respect to their predictions on the same examples.

After optimisation of Equation 3, a subject k can classified according to the following strategies [31]:

- Using only MR data: $\hat{y}_k = \boldsymbol{\alpha_x}^T \boldsymbol{g}_{x,k}$
- Using only US data: $\hat{y}_k = \boldsymbol{\alpha_z}^T \boldsymbol{g}_{z,k}$
- Using both MR and US data: $\hat{y}_k = \frac{1}{2} \left(\boldsymbol{\alpha_x}^T \boldsymbol{g}_{x,k} + \boldsymbol{\alpha_z}^T \boldsymbol{g}_{z,k} \right)$

where $\hat{y}_k$ is the predicted label, and $\boldsymbol{g}_{x,k}$ and $\boldsymbol{g}_{z,k}$ are respectively the Gram matrices from views X and Z for subject k

C. Multi-view classification: A regional approach

In this section we extend the global multi-view classifier outlined in the previous section to propose a novel regional multi-view approach. Fig. 4 summarises this regional approach. Recall that the SSM-based medial mesh used in the formation of the motion atlas indicates to which of the 17 AHA regions each mesh vertex belongs to. Based on this, the MR displacements in atlas space $\boldsymbol{x}_{m,n,k}$ were concatenated by AHA segment into row vectors such that for subject k, $\hat{\boldsymbol{x}}_k^{\text{AHA\_}s} \in \mathbb{R}^{1 \times D_s}$, where D_s = (3 × N × $\mathrm{M_s}$) with N the number of cardiac phases and M_s the number of points in the atlas medial surface mesh in AHA segment s. The row vectors $\hat{\boldsymbol{x}}_k^{\text{AHA\_}s}$ for each subject were stacked to produce matrices $X_{\text{AHA\_}s} = [(\hat{\boldsymbol{x}}_1^{\text{AHA\_}s})^{\mathrm{T}}, ..., (\hat{\boldsymbol{x}}_K^{\text{AHA\_}s})^{\mathrm{T}}] \in \mathbb{R}^{K \times D_s}$, where K is the number of subjects. Likewise, the US displacements $\boldsymbol{z}_{m,n,k}$ were concatenated to form matrices $Z_{\text{AHA\_}s} = [(\hat{\boldsymbol{z}}_1^{\text{AHA\_}s})^{\mathrm{T}}, ..., (\hat{\boldsymbol{z}}_K^{\text{AHA\_}s})^{\mathrm{T}}] \in \mathbb{R}^{K \times D_s}$.

Next, for each AHA segment s we applied the two multi-view classifiers independently: MLDA and MvLapSVM. As a result, for each AHA region the multi-view classifiers predict a label $y_{pred,k}^s$. A

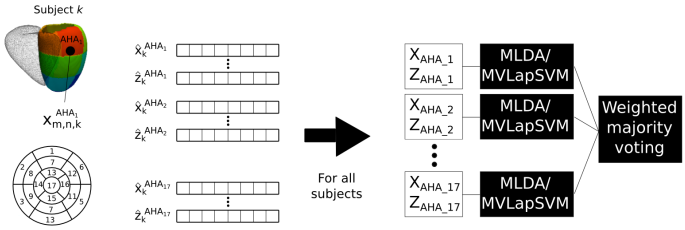

Figure 4: Overview of the regional multi-view learning approach. For each modality, the displacements for each subject k are split into 17 vectors, one for each AHA region. These vectors are used as input to a multi-view learning algorithm, and weighted majority voting is used to decide on the overall classification.

weighted majority voting strategy was used to combine the regional results of each classifier. Details of the calculation of the weights are provided in Section IV-C.

IV. Experiments and Results

Two sets of experiments were performed. The first set of experiments aimed to validate the global multi-view learning approach proposed in Section III-B for identifying patients with DCM, while the second set of experiments aimed at validating the proposed regional approach detailed in Section III-C on the same task.

All experiments were carried out using the Python programming language, using standard Python libraries (Numpy, SciPy, etc.), VTK libraries, and the scikit-learn Python toolkit [33]. Section IV-A details the error measures used for the validation of the proposed algorithms, Section IV-B describes the first set of experiments and presents the results, and Section IV-C presents the validation of the regional approach.

A. Cross validation and evaluation metrics

An 8-fold repeated stratified cross-validation (RSCV) with 100 repetitions was used to compare the performances of the proposed multi-view approaches. In each fold, 44 healthy volunteers and 17 patients were used as training and 2 patients and 6 healthy volunteers as test. As a pre-processing step, the intrinsic dimensionality of both matrices X and Z was estimated using principal component analysis (PCA). The number of PCA modes, L, that capture 95% of the data variance was computed. Then, the first L modes were used when evaluating the PLS and PCA algorithms (see Sections IV-B and IV-C). In each fold, to evaluate the performance of the classification algorithms we computed the confusion matrix, which contains the true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) ratio. As the classes are unbalanced, we computed the balanced accuracy (i.e. the average of the accuracies obtained for each class individually – $BACC = \frac{1}{2}\left(\frac{TP}{P} + \frac{TN}{N}\right)$, as well as the sensitivity (the proportion of patients with DCM correctly classified – $SEN = \frac{TP}{TP+FN}$) and the specificity (the proportion of healthy subjects correctly classified – $SPE = \frac{TN}{TN+FP}$). Finally, we computed the average balanced accuracies, sensitivities and specificities as well as their standard deviations over all folds.

B. Evaluation of global multi-view learning methods

This experiment aims to validate the proposed one-step global multi-view approach for classification. We compared our proposed approach with two classes of comparative method: single modality methods and multiple modality two-step methods. We summarise our proposed approach and the comparative methods in Fig. 5.

The aim of comparing with single modality methods is to evaluate the impact of the multimodal atlas in the classification. To this end, we compared our method to the use of PCA followed by either LDA or support vector machine (SVM) with a radial basis function (RBF) kernel classifier (see Fig. 5). We consider as a baseline the PCA technique trained using US motion data only, as it represents the current state-of-the-art in the use of US data alone for statistical analysis of motion. Since MR is considered to be the gold standard for analysis of cardiac function, we consider the PCA technique trained using only MR data as a reference techniques, regardless of cost or other considerations.

The aim of comparing with the multiple modality two-step approach is to evaluate the impact of our proposed one-step multi-view learning approach over the obvious two-step approach of embedding followed by classification. Previously, in Puyol-Antón *et al.* [3] we showed that partial least squares (PLS) [34] was the optimal multi-view dimensionality reduction algorithm to form a common space between MR and US. Therefore, we applied PLS to reduce the dimensionality of the multimodal data followed by classification using LDA for the linear case and SVM with RBF kernel for the non-linear case.

Note that preliminary results for the linear one-step and two-step approaches were reported in Puyol-Antón *et al.* [24].

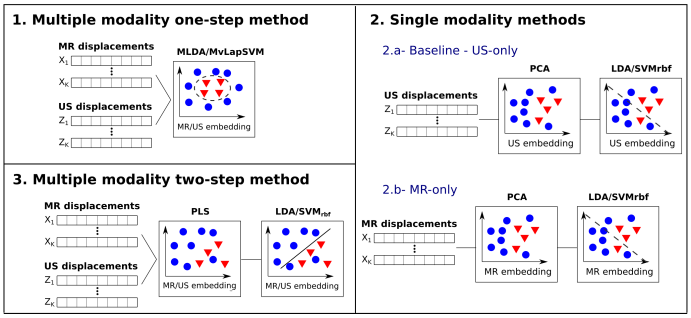

Figure 5: Different approaches for classification: 1. Proposed method (multiple modality one-step method). Comparative approaches: 2. Single modality methods (2.a US-only pipeline, 2.b MR-only pipeline); 3. Multiple modality two-step method.

Table II shows the results of this experiment. For the MLDA classifier the regularisation parameter γ was optimised across values [1, 5, 10, 15, 20] [15]. The RSCV was performed for each of these values of γ, and based on the results, the optimal value of γ was chosen. Similarly, for the MvLapSVM algorithm an RBF kernel was used, and the σ parameter was optimised across values [1E-4, 1E-3, 1E-2, 1E-1, 1, 1E2] and $\gamma_1, \gamma_2, \gamma_3$ were optimised across values [1E-10, 1E-6, 1E-4, 1E-2, 1, 1E2, 1E3]. The RSCV was performed for each of these values of $\gamma_1, \gamma_2, \gamma_3$ and σ, and based on the results, the optimal values were chosen. Note that the test database was used for optimising the parameters and also for the evaluation of the different algorithms due to the limited size of the database. Furthermore, we assess the statistical significance of the accuracy using the proposed method compared to the performance of PCA trained on US data (i.e. the baseline technique) using a Student's t-test (99% confidence).

The results in Table II show that in both the linear and non-linear cases, the use of multi-view algorithms (row 1 - Multiple modality one-step method) results in a statistically significant increase in the accuracy. Both multi-view classifiers have the highest accuracy compared to the two-step approach (row 3 - Multiple modality two-step method). The highest accuracy overall is achieved using MvLapSVM, the non-linear multiple modality one-step method.

Table II: US-based classification balanced accuracy (BACC), sensitivity (SEN) and specificity (SPE) of the proposed and comparative methods and Student's t-test (99% confidence) results. For each group, the first row corresponds to the linear method and the second row to the non-linear method. An asterisk indicates a statistically significant improvement in accuracy over the baseline comparative approach (i.e. 2.a: Single modality US method - PCA + LDA for the linear case and PCA + SVM for the non-linear case). Bold text indicates the method with the highest classification accuracy for the linear and non-linear cases.

Proposed method	**BACC (%)**	**SEN (%)**	**SPE (%)**
1. Multiple modality one-step method			
MLDA	**82.18 (15.0)***	80.50 (26.5)	83.86 (9.9)
MvLapSVM	**92.71 (10.4)***	89.00 (20.8)	95.14 (6.8)
Comparative approaches	**BACC (%)**	**SEN (%)**	**SPE (%)**
2.a. Single modality US method			
PCA_{US} + LDA	74.79 (15.8)	71.05 (28.6)	78.57 (10.1)
PCA_{US} + SVM_{rbf}	87.32 (12.9)	84.50 (23.2)	90.14 (6.6)
2.b. Single modality MR method:			
PCA_{MR} + LDA	84.21 (15.4)*	74.00 (28.8)	90.43 (6.8)
PCA_{MR} + SVM_{rbf}	90.89 (11.7)*	86.50 (22.3)	95.29 (6.7)
3. Multiple modality two-step method			
PLS + LDA	80.07 (16.8)*	75.51 (27.1)	81.86 (9.5)
PLS + SVM_{rbf}	90.39 (12.1)*	87.50 (21.8)	90.57 (10.9)

Using the MvLapSVM algorithm we achieve similar accuracy, sensitivity and specificity to the use of only MR data.

Furthermore, note that the proposed multi-view algorithms (i.e. one-step and two-step, linear and non-linear) are capable of performing classification using data from either view, or both views. Table II focuses on only using US as the input data, while Table III shows the results for the combination of both MR and US data.

Table III: MR and US classification balanced accuracy, sensitivity and specificity of the proposed and comparative methods. For each group, the first row corresponds to the linear method and the second row to the non-linear method. Bold text indicates the method with the highest classification accuracy for the linear and non-linear cases.

Proposed method	**BACC (%)**	**SEN (%)**	**SPE (%)**
1. Multiple modality one-step method			
MLDA	**88.93 (11.4)**	88.00 (21.5)	89.22 (8.1)
MvLapSVM	**94.86 (9.5)**	91.00 (19.3)	96.57 (6.1)
Comparative approaches	**BACC (%)**	**SEN (%)**	**SPE (%)**
3. Multiple modality two-step method			
PLS + LDA	77.25 (16.9)	72.50 (29.6)	82.00 (10.1)
PLS + SVM_{rbf}	93.79 (9.7)	94.00 (16.3)	95.71 (6.6)

Comparing Tables II and III, we can see that the multi-view classifier algorithms have a higher accuracy when using both views as input data. We can also see that, for the comparative multiple modality two-step method, accuracy in the linear case is slightly lower, while in the non-linear case the accuracy is slightly higher. This can be explained by the observation that non-linear classifiers can learn the non-linear structure of the data which is likely to result in improved classification performance. It is likely that multi-view machine learning algorithms provided the optimal performances when using the two views as they were better able to exploit the inter-modality relationship when performing classification. This section showed that the use of only one view (i.e. US) only slightly reduced the accuracy of the proposed algorithms when compared to the use of both views (i.e. MR and US). In both cases, MLDA and MvLSVM, the prediction function (Eqs. 1 and 3) is the average of the prediction functions from the two views. In the ideal case, the embedding space for MR and US should be the same, but as the data might have some noise and the number of training subjects is limited, this embedding might vary as reported in Puyol-Antón *et al.* [3]. For this reason, it is probable that the use of the two views provided a more robust embedding space.

Based on these results, we conclude that the highest accuracy is achieved using our proposed multiple modality one-step multi-view classifiers, whether using only US data or both MR and US data as input.

C. Evaluation of regional multi-view learning methods

This section reports classification results for the regional approach described in Section III-C. Table IV shows the results of the proposed regional approach compared to the global approach. A Student's t-test (99% confidence) was used to evaluate the statistical significance of the accuracy results between global and regional methods.

Similar to the global approach, but now for each AHA region, the regularisation parameter γ for MLDA was optimised across values [1, 5, 10, 15, 20], and for the MvLapSVM algorithm the σ parameter was optimised across values [1E-4, 1E-3, 1E-2, 1E-1, 1, 1E2] and $\gamma_1, \gamma_2, \gamma_3$ were optimised across values [1E-10, 1E-6, 1E-4, 1E-2, 1, 1E2, 1E3]. Once the parameters for the linear and non-linear method were optimised per AHA region, the predicted labels were stored. Then, they were combined using a weighted majority voting strategy. The optimal weights for each AHA segment were optimised using a randomised search on hyper parameters algorithm [35]. More specifically, the weights for each AHA segment were sampled from a uniform distribution in the interval [0,1], then the sampled weights were used to combine the predicted labels and the accuracy, sensitivity and specificity were computed. This process was iteratively repeated 500 times and the best optimal weights were selected according to the highest accuracy.

Table IV: US-based classification balanced accuracy, sensitivity and specificity of global and regional approaches and Student's t-test (99% confidence) results. An asterisk indicates a statistically significant improvement in accuracy over the global approach.

Global Methods	**BACC (%)**	**SEN (%)**	**SPE (%)**
MLDA	82.18 (15.0)	80.50 (26.5)	83.86 (9.9)
MvLapSVM	92.71 (10.4)	89.00 (20.8)	95.14 (6.8)
Regional Methods	**BACC (%)**	**SEN (%)**	**SPE (%)**
MLDA	87.71 (12.6)*	85.00 (23.1)	90.43 (6.7)
MvLapSVM	94.32 (11.1)*	93.00 (17.5)	96.57 (6.2)

The results in Table IV show that the regional method outperforms the global method in both linear and non-linear cases, with both differences being statistically significant.

Figure 6 shows the estimated weights for the linear and non-linear cases. These show that the basal and mid segments of the anterior wall are the areas with the highest impact for classification. This is consistent with the findings of previous work that reported large motion in the basal and mid free wall segments [36] and the importance of these areas for cardiac function [37]. US based assessment of the motion of the basal anterior wall of the myocardium is known to be subject to uncertainty [38], due to limited acoustic windows and significant noise in the far field of the US beam. This could explain the improvement seen in our method, as inclusion of MR-derived information in the multimodal motion atlas adds complementary information.

Similar to the previous section, the accuracy of the different classifiers using both MR and US data was also computed. Table V shows the results of the proposed regional algorithms using both modalities. A similar pattern to that in the previous section can be observed. The use of both modalities provides a richer motion description and therefore the classifiers achieve a higher accuracy.

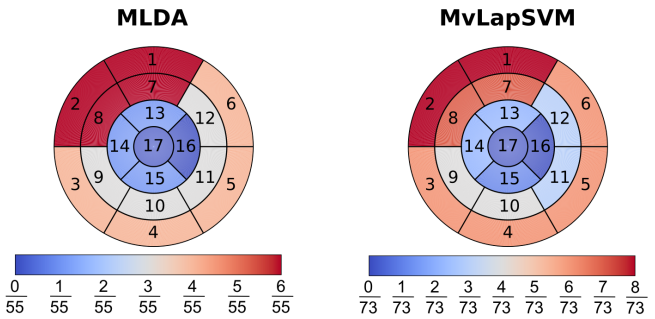

Figure 6: On the left estimated weights for MLDA, and on the right estimated weights for MvLapSVM using the randomised search on hyper parameters algorithm. The colour bar shows the ranges of values for the estimated weights.

However, in this case there is only a statistically significant increase in accuracy for the linear method compared to the global method.

Table V: MR and US classification balanced accuracy, sensitivity and specificity of global and regional approaches and Student's t-test (99% confidence) results. An asterisk indicates a statistically significant improvement in accuracy over the global approach.

Global Methods	**BACC (%)**	**SEN (%)**	**SPE (%)**
MLDA	88.93 (11.4)	88.00 (21.5)	89.22 (8.1)
MvLapSVM	94.86 (9.5)	91.00 (19.3)	96.57 (6.1)
Regional Methods	**BACC (%)**	**SEN (%)**	**SPE (%)**
MLDA	91.25 (10.5)*	91.50 (18.9)	91.00 (6.9)
MvLapSVM	94.96 (10.2)	90.50 (19.7)	97.43 (5.5)

D. Clinical validation

As described at the beginning of this section, to validate the proposed algorithms we used an 8-fold RSCV with 100 repetitions. In addition to computing the balanced accuracy, sensitivity and specificity for each fold, we also recorded how many times each patient was misclassified. In this section we investigate whether a relationship exists between the subjects who were frequently misclassified and clinical parameters used to detect patients suffering from DCM. As described in Section I-C, DCM diagnosis criteria include reduced global EF, ventricular dilatation, as well as impaired systolic function.

Fig. 7 reports the relationship between EF and the misclassification rate, which we defined as the number of times a subject was misclassified over the 100 repetitions divided by the total number of classifications for that subject. For both global and regional methods, we can see that the misclassification rate for DCM patients increases with higher EF, which is consistent with the clinical perspective that EF decreases with the progression of DCM, as the LV dilates and the cardiac function decreases. The EFs for the most misclassified DCM patients are typically over 50%, which is in the same range as the volunteer EFs reported in Table I. Similarly, the EFs for the most misclassified healthy subjects are closer to the EFs for DCM patients.

As a final experiment we determined the accuracy of detecting DCM patients only using EF values, which is considered the primary diagnosis criterion to identify patients suffering DCM. In this case, we used values reported in Table I of EF for each subject to train a LDA classifier. Similar to the previous experiments, we computed the balanced accuracy, sensitivity and specificity.

The results in Table VI show that the proposed method achieves higher accuracy than only using EF values. Furthermore, the EF based

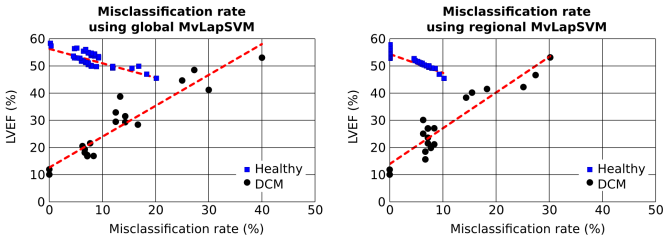

Figure 7: Relationship between EF (%) and misclassification rate for the preferred method (MvLapSVM) for the global approach (left figure) and the regional approach (right figure) for DCM patients (black circles) and healthy volunteers (blue squares).

Table VI: Comparison between EF classifier and regional non-linear multi-view classifier: balanced accuracy, sensitivity and specificity and Student's t-test (99% confidence) results. An asterisk indicates a statistically significant improvement in accuracy over the global approach.

Method	**BACC (%)**	**SEN (%)**	**SPE (%)**
LDA for EF	82.5 (17.1)	100 (0)	65.0 (34.2)
MvLapSVM	94.32 (11.1)*	93.00 (17.5)	96.57 (6.2)

classifier always correctly classifies healthy patients (the sensitivity), but is less successful in classify DCM patients (specificity). Therefore, we believe that the use of the atlas and more localised motion descriptors helps in the identification of DCM patients.

V. Discussion and Conclusion

Previously in Puyol-Antón *et al.* [39] we described the construction and evaluation of a multimodal cardiac motion atlas that related MWMA parameters derived from MR and US data. In this paper, we have proposed a novel extension to this atlas and demonstrated its application to the task of identifying DCM patients using only US data but at the same time taking advantage of the implicit relationship between MR and US learned from the atlas. This was made possible by the inclusion of multi-view learning algorithms in which MR and US were considered as different views of the same underlying phenomenon (i.e. cardiac mechanics). More specifically, two multi-view classifiers were evaluated (MLDA and MvLapSVM), and both were implemented using global and regional approaches. To the best of our knowledge this is the first time that multi-view learning was applied to cardiac motion data.

Our application in this paper was the identification of DCM, but the same pipeline could be used to identify other cardiovascular diseases. We believe that this approach has great clinical utility, as 3D tagged MR data is considered to be the gold standard for estimation of MWMA parameters due to its high contrast, high resolution and good spatial coverage. However, the use of MR scanners is time consuming, expensive and 3D tagged MR sequences are not widely available. Furthermore, its use is contraindicated in patients with cardiac pacemakers or metallic cardiovascular electronic devices. Our pipeline based only on US data opens up the possibility of accurately characterising cardiac function and anatomy using a low cost, more accessible and portable imaging modality.

Our results showed that the use of multi-view machine learning algorithms resulted in a statistically significant increase in classification accuracy compared to the use of only US data without the multimodal atlas. Furthermore, the comparison of the global and regional approaches showed that the regional approach resulted in a statistically significant improvement in terms of accuracy. The highest accuracy was 94.32% and it was achieved using regional

MvLapSVM. We believe that the better performance of the regional approach is because of the high dimensionally of the input data and the limited size of the training database. The regional approach was able to find more meaningful descriptors from the lower dimensional regional data to detect patients suffering from DCM.

A clear indicator for detecting DCM is the size of the LV. However, in our pipeline we aimed to use only motion-based information to detect this group of patients without including shape information. We used DCM as an exemplar application to demonstrate the power of our method. In reality, it is likely that inclusion of morphological information would be beneficial to the identification of DCM or other patient groups, and this will be the subject of future work.

Future work will also focus on incorporating deformation parameters (strain, strain rate) and non-imaging data (e.g. age, gender and information about clinical history) to have a more complete diagnostic pipeline. Furthermore, we would like to test the same pipeline in other groups of patients, and try to create a multi-class classifier to be able to detect different diseases.

VI. Conclusion

In conclusion, we believe that the work we have presented represents an important contribution to the understanding of cardiac motion. The intended workflow of the developed pipeline is to make use of the prior knowledge from the multimodal atlas to enable robust extraction of indicators from 3D US images for detecting DCM patients. In a clinical setting, we would envisage a scenario in which our pipeline would be embedded into the US scanner to provide real-time information to the sonographer during acquisition.

References

[1] M. D. Cerqueira *et al.*, "Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart," *Circulation*, vol. 105, no. 4, pp. 539–542, 2002.

[2] R. Jasaityte, B. Heyde, and J. Dhooge, "Current state of three-dimensional myocardial strain estimation using echocardiography," *Journal of the American Society of Echocardiography*, vol. 26, no. 1, pp. 15–28, 2013.

[3] E. Puyol-Antón *et al.*, "A multimodal spatiotemporal cardiac motion atlas from MR and ultrasound data," *Medical Image Analysis*, 2017.

[4] D. Perperidis, R. Mohiaddin, and D. Rueckert, "Construction of a 4D statistical atlas of the cardiac anatomy and its use in classification," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*. Springer, 2005, pp. 402–410.

[5] M. Sinclair *et al.*, "Myocardial strain computed at multiple spatial scales from tagged magnetic resonance imaging: Estimating cardiac biomarkers for CRT patients," *Medical Image Analysis*, 2017.

[6] D. Peressutti *et al.*, "A framework for combining a motion atlas with non-motion information to learn clinically useful biomarkers: Application to cardiac resynchronisation therapy response prediction," *Medical image analysis*, vol. 35, pp. 669–684, 2017.

[7] W. Bai *et al.*, "A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion," *Medical image analysis*, vol. 26, no. 1, pp. 133–145, 2015.

[8] N. Duchateau *et al.*, "A spatiotemporal statistical atlas of motion for the quantification of abnormal myocardial tissue velocities," *Medical image analysis*, vol. 15, no. 3, pp. 316–328, 2011.

[9] ——, "Constrained manifold learning for the characterization of pathological deviations from normality," *Medical image analysis*, vol. 16, no. 8, pp. 1532–1549, 2012.

[10] C. Hoogendoorn *et al.*, "A high-resolution atlas and statistical model of the human heart from multislice CT," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 1, pp. 28–44, 2013.

[11] K. Punithakumar *et al.*, "Regional heart motion abnormality detection: An information theoretic approach," *Medical image analysis*, vol. 17, no. 3, pp. 311–324, 2013.

[12] J. Zhao *et al.*, "Multi-view learning overview: Recent progress and new challenges," *Information Fusion*, vol. 38, pp. 43–54, 2017.

[13] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.

[14] F. Bach and M. Jordan, "Kernel independent component analysis," *Journal of machine learning research*, vol. 3, no. Jul, pp. 1–48, 2002.

[15] S. Sun, X. Xie, and M. Yang, "Multiview uncorrelated discriminant analysis," *IEEE transactions on cybernetics*, vol. 46, no. 12, pp. 3272–3284, 2016.

[16] M. Kan *et al.*, "Multi-view discriminant analysis," *Computer Vision–ECCV 2012*, pp. 808–821, 2012.

[17] M. Gönen and E. Alpaydın, "Multiple kernel learning algorithms," *Journal of machine learning research*, vol. 12, no. Jul, pp. 2211–2268, 2011.

[18] S. Sanchez-Martinez *et al.*, "Characterization of myocardial motion patterns by unsupervised multiple kernel learning," *Medical image analysis*, vol. 35, pp. 70–82, 2017.

[19] M. Marciniak *et al.*, "A multiple kernel learning framework to investigate the relationship between ventricular fibrillation and first myocardial infarction," in *International Conference on Functional Imaging and Modeling of the Heart*. Springer, 2017, pp. 161–171.

[20] C. Yancy *et al.*, "2013 ACCF/AHA guideline for the management of heart failure," *Circulation*, pp. CIR–0b013e31 829e8776, 2013.

[21] H. HFSA, "HFSA 2010 comprehensive heart failure practice guideline," pp. e1–e194, 2010.

[22] M. Merlo *et al.*, *Dilated Cardiomyopathy: Usefulness of Imaging in Prognostic Stratification and Choice of Treatment*. Cham: Springer International Publishing, 2014, pp. 75–81.

[23] B. Bozkurt and D. Mann, "Dilated cardiomyopathy," in *Cardiovascular Medicine*. Springer, 2007, pp. 1233–1259.

[24] E. Puyol-Antón *et al.*, "Multiview machine learning using an atlas of cardiac cycle motion," in *International Workshop on Statistical Atlases and Computational Models of the Heart*, 2017, pp. 1867–1876.

[25] C. Tobon-Gomez *et al.*, "Benchmarking framework for myocardial tracking and deformation algorithms: An open access database," *Medical Image Analysis*, vol. 17, no. 6, pp. 632–648, 2013.

[26] M. Sinclair *et al.*, "Fully automated segmentation-based respiratory motion correction of multiplanar cardiac magnetic resonance images for large-scale datasets," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 332–340.

[27] J. Gower and G. Dijksterhuis, *Procrustes problems*. Oxford University Press Oxford, 2004, vol. 3.

[28] D. Rueckert *et al.*, "Nonrigid registration using free-form deformations: application to breast MR images," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 8, pp. 712–721, 1999.

[29] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal processing magazine*, vol. 16, no. 6, pp. 22–38, 1999.

[30] L. Tu, "An introduction to manifolds, ch. 14," 2007.

[31] S. Sun, "Multi-view laplacian support vector machines," in *International Conference on Advanced Data Mining and Applications*. Springer, 2011, pp. 209–222.

[32] S. Melacci and M. Belkin, "Laplacian support vector machines trained in the primal," *Journal of Machine Learning Research*, vol. 12, no. Mar, pp. 1149–1184, 2011.

[33] S.-l. D. Scikit, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[34] S. Wold *et al.*, "Multivariate data analysis in chemistry," in *Chemometrics*. Springer, 1984, pp. 17–95.

[35] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, no. Feb, pp. 281–305, 2012.

[36] S. Kleijn *et al.*, "Normal reference values of left ventricular strain using three-dimensional speckle tracking echocardiography: results from a multicentre study," *European Heart Journal-Cardiovascular Imaging*, vol. 16, no. 4, pp. 410–416, 2014.

[37] M. Carlsson *et al.*, "Atrioventricular plane displacement is the major contributor to left ventricular pumping in healthy adults, athletes, and patients with dilated cardiomyopathy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 292, no. 3, pp. H1452–H1459, 2007.

[38] W. Armstrong and T. Ryan, *Feigenbaum's echocardiography*. Lippincott Williams & Wilkins, 2012.

[39] E. Puyol-Antón *et al.*, "Towards a multimodal cardiac motion atlas," in *Biomedical Imaging (ISBI), 2016 13th IEEE International Symposium on*, 2016.